DOCUMENT RESUME

ED 447 234 UD 033 853

AUTHOR Shaul, Marnie S.
TITLE Education for Disadvantaged Children: Research Purpose and Design Features Affect Conclusions Drawn from Key Studies. Report to Congressional Requesters.
INSTITUTION General Accounting Office, Washington, DC. Health, Education, and Human Services Div.
REPORT NO GAO/HEHS-00-168
PUB DATE 2000-08-00
NOTE 28p.
AVAILABLE FROM U.S. General Accounting Office, P.O. Box 37050, Washington, DC 20013. Tel: 202-512-6000; Fax: 202-512-6061; e-mail: info@www.gao.gov; Web site: http://www.gao.gov.
PUB TYPE Reports - Evaluative (142)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Compensatory Education; Elementary Secondary Education; Federal Aid; Federal Legislation; *Federal Programs; *Longitudinal Studies; *Research Design; Research Methodology
IDENTIFIERS *Elementary Secondary Education Act Title I

ABSTRACT

To inform reauthorizations of Title I, the federal government's largest program for elementary and secondary education, the Congress has required the Department of Education to conduct national assessments of Title I. Two of these studies have gathered Title I data over several years: the Prospects study, completed in 1997; and the ongoing Longitudinal Evaluation of School Change and Performance (LESCP), which is due to be completed in 2001. Congress asked the General Accounting Office to compare these two studies and analyze their strengths and weaknesses. Documents produced in each study and interviews with project staff and other experts were used to compare the two studies. There was a consensus that the purpose of the Prospects study was to assess the effectiveness of Title I, but there was less agreement about the precise purpose of the LESCP. This lack of agreement about purpose made it hard to predict the degree to which the LESCP report will meet the needs of Congress, other policymakers, and educators. The two studies used similar processes for contracting out data collection and analysis, obtaining comments from review panels, and issuing results. For both studies, major design features influenced the study's ability to address its overall design. The large sample for the Prospects study supported fairly strong conclusions about the effectiveness of Title I. The depth of focus provided by the longitudinal design of the LESCP is likely to be a key strength, but the study uses a smaller, nonrepresentative sample, a design limitation that may limit the usefulness of results. Both studies have provided valuable information, but no single study can fulfill the diverse information needs of the Congress, the Department of Education, all of the states, and educators. One appendix contains a chart comparing the two studies, and the other contains comments from the Department of Education. (SLD)

Reproductions supplied by EDRS are the best that can be made from the original document.

UD

United States General Accounting Office

GAO

Report to Congressional Requesters

August 2000

EDUCATION FOR DISADVANTAGED CHILDREN

Research Purpose and Design Features Affect Conclusions Drawn From Key Studies

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

ED 447 234

UD033853

GAO/HEHS-00-168

BEST COPY AVAILABLE

Contents

Letter 3

Appendixes

Appendix I: Comparison of Prospects and LESCP Studies 22

Appendix II: Comments From the Department of Education 25

Abbreviations

IRP	Independent Review Panel
LESCP	Longitudinal Evaluation of School Change and Performance
RFP	request for proposals
TWG	Technical Working Group

GAO
Accountability * Integrity * Reliability

United States General Accounting Office
Washington, D.C. 20548

Health, Education, and
Human Services Division

B-285997

August 31, 2000

The Honorable William F. Goodling
Chairman, Committee on Education and the Workforce
House of Representatives

The Honorable Michael N. Castle
Chairman, Subcommittee on Early Childhood, Youth
and Families
Committee on Education and the Workforce
House of Representatives

The Honorable Peter Hoekstra
Chairman, Subcommittee on Oversight
and Investigations
Committee on Education and the Workforce
House of Representatives

Title I, the federal government's largest program for elementary and secondary education, is primarily directed at assisting disadvantaged children in high-poverty schools. To inform the 1994 and impending reauthorizations of this important program,[1] the Congress has required the Department of Education to conduct a national assessment of Title I—a broad study to examine the overall progress of students served by the program and the implementation of its key provisions. Education has conducted a number of research studies to support these national assessments. Two of these studies gathered data on Title I students over several years—the Prospects study, completed in 1997, and the ongoing Longitudinal Evaluation of School Change and Performance (LESCP),

[1]Title I has been scheduled for reauthorization every 5 years, but the 1999 reauthorization has not yet been completed.

which is due to be completed in 2001.[2] Recently, concerns have been raised about the quality and usefulness of the LESCP study. These concerns have focused on the quality of the study design, whether Education obtained sufficient input from external experts, and the extent to which Education provided data from the study to inform the Congress's debate over reauthorization in a timely manner. You asked us to review and compare LESCP to its predecessor, the Prospects study. In this report, we (1) review the purpose of each study and how these purposes relate to the needs of policymakers and educators; (2) describe the process used to design and implement the studies, including obtaining feedback from review panels and releasing the results to the Congress and the public; and (3) analyze the studies' strengths and limitations in light of their purposes and determine the effect these strengths and limitations have on the conclusions that can be drawn from the data.

We reviewed the interim and final reports from the Prospects study, the interim report from the LESCP study, and the final reports of the 1993 National Assessment of Chapter 1 and the 1999 National Assessment of Title I.[3] We also reviewed internal Education documents relating to the design and analysis of each study. We analyzed the minutes of all the meetings of the current Independent Review Panel (IRP), a congressionally mandated advisory panel to Education, and reviewed the panel's final report to Congress. We also reviewed the minutes of all the meetings of the LESCP Technical Working Group (TWG), a panel of expert researchers convened by Education to advise the Department and its contractors on the study. Finally, we interviewed Education staff, including Title I program staff and staff of the Planning and Evaluation Service, which directed both studies; IRP and TWG members; the project directors for the LESCP and Prospects studies from the primary contracting organizations; and outside

[2]Abt Associates, Inc., *Prospects: Final Report on Student Outcomes*, prepared under contract to the Planning and Evaluation Service, U.S. Department of Education (Washington, D.C.: U.S. Department of Education, Apr. 1997); Abt Associates, Inc., *Prospects: The Congressionally Mandated Study of Educational Growth and Opportunity, Interim Report*, prepared under contract to the Planning and Evaluation Service, U.S. Department of Education (Washington, D.C.: U.S. Department of Education, July 1993); Policy Studies Associates, Inc. and Westat, Inc., *The Longitudinal Evaluation of School Change and Performance in Title I Schools, Interim Report*, prepared under contract to the Planning and Evaluation Service, U.S. Department of Education (Washington, D.C.: U.S. Department of Education, July 1999).

[3]Before the 1994 reauthorization, the Title I program was called "Chapter 1." Throughout this report, we refer to the program as Title I.

experts. We performed our work between June and August 2000 in accordance with generally accepted government auditing standards.

Results in Brief

While Education staff, contractors, and members of advisory panels all agreed that the purpose of the Prospects study was to assess the effectiveness of Title I, considerably less agreement existed on the precise purpose of the LESCP study. These differences centered primarily around how much emphasis the study would place on the Title I program. For example, although some expected LESCP to serve as another vehicle for evaluating Title I's overall effectiveness, other individuals expected LESCP to evaluate specific education reform policies implemented under the 1994 reauthorization of Title I. Still other individuals expected the LESCP study to focus primarily on changes in instructional practice, with Title I issues treated more as a contextual factor than as a major focus of the analysis. This lack of agreement about LESCP's purpose created unclear expectations for the study, making it difficult to predict the degree to which the final report will meet the needs of the Congress, other policymakers, and educators.

The LESCP and Prospects studies used similar processes for contracting out data collection and analysis, obtaining comments from review panels, and releasing the results. Both studies were conducted by outside research organizations under contracts with Education. For both Prospects and LESCP, Education and the contractors obtained advice from two panels of experts—a congressionally mandated review panel that advised Education on policy issues and a separate technical panel that provided feedback specific to the individual study. The panels that advised the Prospects study had proportionally greater representation from educators, while LESCP's panels have proportionally more researchers. For both the LESCP and Prospects studies, the 3 to 4 year longitudinal data collection and the complexity of the implementation process posed challenges for providing study results in time to meet Congress's 5-year reauthorization schedule. To provide information for reauthorization, Education issued interim reports to Congress in both cases.

For both Prospects and LESCP, major design features influenced the study's ability to address its overall purpose. The design of each study has both strengths and potential limitations. For example, Prospects' large, representative sample supported fairly strong conclusions about the effect of Title I, although the study design did not allow researchers to definitively measure how students would have performed if they did not have access to

Title I services. The LESCP study is not yet complete, but the depth of information provided by its longitudinal focus is likely to be a key strength. However, LESCP, which uses a smaller, nonrepresentative sample, suffers from design limitations that will restrict its ability to fully satisfy any of the three potential purposes—evaluation of Title I, standards-based reform, or instructional practices—envisioned by Education, contractors, and panel members. Because several design limitations are directly related to measuring the effects of Title I and standards-based reform, they will likely have a major effect on the conclusions that could be drawn in these areas of education policy. They will, however, have a less serious effect on the conclusions that could be drawn about instructional practices.

Background

Title I of the Elementary and Secondary Education Act was established in 1965 to help schools meet the needs of economically and educationally disadvantaged students. Title I's $7.9 billion budget is targeted primarily—but not exclusively—to schools with a relatively high percentage of students from low-income families. Although it is the federal government's largest elementary and secondary education program, Title I accounts for a small share of total education expenditures—about 3 percent. However, for many poor schools Title I is a key source of funding for items such as supplementary instruction, professional development, new computers, and after-school or other extended-time programs.

In 1994, the Congress made several significant changes to the program. First, the Congress changed the rules for allocating Title I dollars in an effort to direct more funding to the neediest schools. Second, in the 1994 reauthorization the Congress increased the number of schools eligible to use Title I funds on a schoolwide basis. Title I has traditionally expected schools to direct program funds to students who are low-achieving or at highest risk for school failure (called targeted assistance). However, a provision known as the schoolwide program allows a school to spend its Title I funds to improve the school as a whole, rather than targeting Title I dollars to low-achieving students. Because all students in the school can benefit, under a schoolwide program it is more difficult to distinguish the effect of Title I services from the effect of a school's overall instructional program. The number of schools adopting the schoolwide approach has increased dramatically, to nearly half of all Title I schools, since the 1994 reauthorization expanded eligibility for schoolwide status.

In the 1994 legislation, the Congress also established a new reform policy, commonly called standards-based reform. Under these new requirements,

states (in addition to their general oversight responsibilities) were to collect and report information on educational outcomes and to hold schools and districts accountable for results. Specifically, states were required to develop content standards to describe what students need to know and performance standards to describe their expected proficiency in at least the core subject areas of reading or language arts and mathematics. States were then to establish assessments to measure how students are doing in relation to these standards. Each state must report the results of assessments. In addition to the overall results, states must collect and publicly report assessment results by six specified student categories—gender, racial and ethnic group, English proficiency status, disability status, migrant status, and economic status. States must also develop criteria for determining whether schools and districts are performing satisfactorily.[4] These standards-based reform requirements were designed to increase Title I's focus on educational outcomes for all students, not just those children who had been traditionally targeted under the program.

To support the national assessments of Title I, Education issued contracts to outside researchers to conduct several major projects. For the earlier reauthorization, Education's major Title I research effort was the Prospects study, a large longitudinal study of Title I students and schools. The Prospects study also responded to an additional, more specific congressional mandate that required Education to conduct a longitudinal study to compare students who received Title I services with students who did not. Prospects began gathering student data in 1991; although not completed until 1997, the study played an important role in informing the 1994 reauthorization of Title I. In the more recent national assessment, Education relied on data from several studies. One of these studies, LESCP, is also a longitudinal study of Title I schools. Although the legislative mandate for the most recent national assessment required longitudinal data, the legislative requirements were less specific than those set out in the mandate that governed the Prospects study.

[4]States had until the 1997-98 school year to develop content and performance standards. Education, as authorized by statute, extended the deadline for performance standards for many states to coincide with the deadline for assessments, which were to be finalized by the 2000-01 school year.

Strong Agreement Existed on Purpose of Prospects Study But Did Not for LESCP

Education staff, panel members, contractor staff, and the interim and final reports from the Prospects study consistently described the purpose of the Prospects study as evaluating the effectiveness of the Title I program. This purpose reflected the specific legislative mandate for the Prospects study. Consistent with this purpose, the main finding of the Prospects study was that students who received Title I services started below the achievement level of their peers and that this initial gap remained essentially unchanged as students moved into higher grades. (See app. I for more information about the findings of the Prospects study.) The Prospects data directly addressed congressional needs for information on overall program effectiveness, and it also provided some useful information for other policymakers and educators.

Considerably less agreement existed among Education staff, panel members, and contractors on the precise purpose of the LESCP study, especially as it relates to the Title I program. LESCP's legislative requirements were less specific than those of the Prospects study, which responded to a specific congressional mandate. Education developed the following study questions for LESCP: (1) To what extent are changes occurring in what is being taught in reading and mathematics in the classrooms in the study? (2) To what extent are changes occurring in how instruction is being delivered? (3) To what extent are students showing changes in performance? and (4) How do recent revisions in Title I contribute to these changes? Despite having these study questions, Education staff, panel members, and contractors—all of whom were closely connected with the LESCP project—differed in their view of the study's purpose. These differences appeared to be primarily concerned with the emphasis the study would place on the fourth study question—the role of Title I. For some staff and panel members, the fourth question was the key research question for the project. For others, the first three questions were central to the study, and Title I was to be treated more as a contextual factor than a major focus of analysis. As a result, individuals seeing the same study questions held very different views on the overall purpose of the project. For example, while several individuals stated that LESCP's purpose was to evaluate the effectiveness of the Title I program, others told us that this was not the study's purpose. Similarly, several people expected LESCP to evaluate standards-based reform, while others stated that evaluating standards-based reform was beyond the study's scope. Yet others expected LESCP to focus primarily on the effect of instructional practices on student achievement.

LESCP's purpose was not fully clarified by the release of an interim report. The report reflected a primary emphasis on studying the effect of instructional practices on student achievement. However, several panel members expected the interim report to include more information on Title I or standards-based reform. The report found that certain instructional practices, such as having students talk in small groups about what they read, were correlated with improved student performance in fourth grade reading. The report also found that there was little change in these instructional practices between 1997 and 1998. (See app. I for more information on the findings of the interim LESCP report.) Education expects some findings related to standards-based reform to be included in the final LESCP report, which is due for release in 2001. In the absence of the final report, however, the divergent views on LESCP's purpose make it difficult to predict whose needs the study will address—the Congress's need for information related to Title I effectiveness, Education's and the states' needs for information on evaluating the importance of various standards-based reform practices, or educators' needs for information on the most effective instructional practices for promoting student achievement.

Studies Used Similar Processes for Contracting, Panel Review, and Release of Data

The process for implementing a longitudinal study—including contracting with outside research organizations and obtaining expert advice on important implementation issues—can have a major effect on the quality and timeliness of the final product. For both Prospects and LESCP, Education emphasized quality by gathering longitudinal data and consulting with expert panels to obtain advice and feedback on important implementation issues. In both cases, however, gathering data over 3 to 4 years and following a multistep contracting process posed challenges for providing information to the Congress in time to meet the Title I reauthorization schedule.

Both Studies Followed Multistep Contracting Process

Both the Prospects and LESCP studies were conducted by outside research organizations under contracts that included data collection, analysis, and report writing. Education used a similar competitive bidding process for issuing these contracts. Under this process, Education staff developed a request for proposals (RFP) that included detailed specifications for each study. For example, the RFP for LESCP specified key design features such as the sample size, characteristics of the sample, data elements (including both state assessments and standardized exam results, for example), and the time frame for data collection. Potential contractors then submitted

proposals to Education. Education staff and peer reviewers evaluated these proposals on the basis of the quality of the proposal, the contractors' experience in education evaluation, and the contractor's bid. After the award, a designated Education staff member supervised the project and monitored contractor progress. For both projects, the process for awarding contracts took several months.

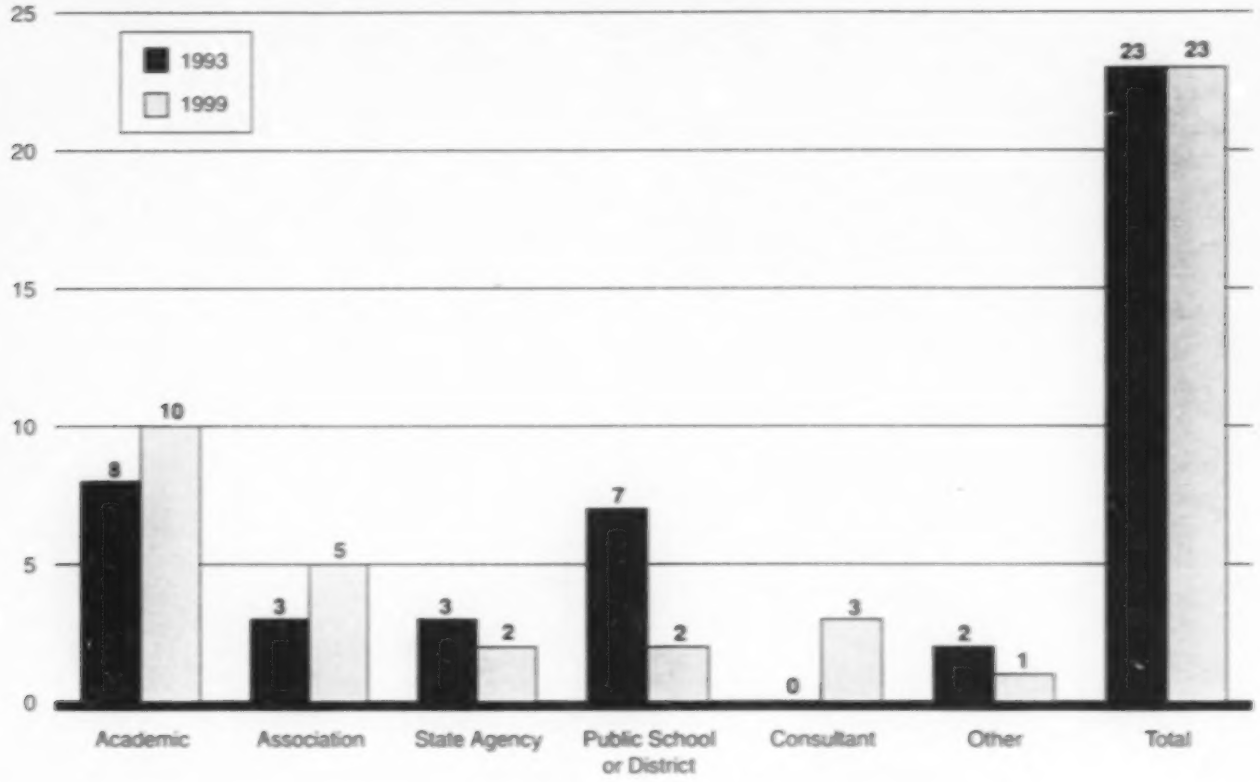
Experts Provided Policy and Technical Advice for Both Prospects and LESCP Studies

For both the Prospects and LESCP studies, Education convened two types of expert panels—one to provide broad policy guidance and another to provide expert advice on the technical aspects of the study. The policy panel—the IRP—advised Education on broad issues that cut across several studies, while the TWG provided advice and guidance on technical issues specific to each study.

For each of the two previous Title I reauthorizations, the Congress required Education to establish an IRP to advise the Department on its research efforts concerning the National Assessment of Title I. In its quarterly meetings, the IRP has provided feedback on the relationship between broad education reform issues and Education's research efforts and has raised issues about the limitations of specific research projects. Each IRP has also chosen to submit its own report to Congress to accompany Education's report on the National Assessment of Title I. These reports provide the panels' perspective and recommendations on important reauthorization issues, such as ensuring high-quality professional development for teachers.

Congress required the IRPs to include researchers, state and local practitioners, and other appropriate individuals. Both panels included representatives from academia, education associations, state and local education agencies, and others. However, the current panel includes a greater percentage of researchers and association representatives, while the 1993 panel included a greater percentage of educators, especially from the local level. Fig. 1 details the affiliations of IRP members in 1993 and 1999.

Figure 1: Primary Affiliation of Independent Review Panel Members, 1993 and 1999



In contrast to the IRPs, the TWGs are designed to assist the Department and the contractors on a single study. Initially, contractors recommend TWG panel members in their proposals, and Education selects the panel once the award has been made. TWG members are primarily researchers, selected to provide subject area and methodological expertise in areas important to the study. The TWG's role is to provide guidance and advice in implementing the study design, analyzing the data, and writing the report. For example, the LESCP TWG reviewed the survey instruments and made suggestions to reduce the length of the proposed teacher questionnaire. However, the TWG members are generally not involved in the early study design because the group is not formed until after many of the major design decisions have already been made. Moreover, TWG input cannot result in major modifications to the study design because the contract, which

includes detailed design specifications, has already been finalized and may need to be re-competed.

TWGs were used on both the Prospects and LESCP studies. The Prospects study also sought advice from a "stakeholders group," which included representatives from state education agencies, local educators, and advocates. However, Education staff told us that the two panels often duplicated efforts, with both groups raising similar issues. For the LESCP study, Education incorporated representatives from state education agencies into the TWG, instead of convening a stakeholders group. Combined with the reduced number of educators on the IRP, the elimination of the stakeholders group resulted in fewer educators—especially at the local level—involved in advising Education on LESCP as compared with the Prospects study. Several individuals involved with the LESCP study stated that additional input from educators, particularly in the early stages of study design, would have been helpful. However, Education staff and some panel members also told us that it was sometimes difficult to get educators to participate in their advisory panels.

For Both Studies, Data Collection Requirements Posed Challenges to Meeting Reauthorization Schedule

Both the Prospects and LESCP studies faced challenges in trying to provide Congress with timely information to inform reauthorization of Title I. Title I was scheduled to be reauthorized every 5 years. However, for each study, more than 5 years were needed to design the study, issue the contract, obtain the required clearance for the study instruments from the Office of Management and Budget, collect and analyze several years' worth of data, and complete the final report. Both Prospects and LESCP provided some information for reauthorization through interim reports. Education also may provide to Congress preliminary information from the studies before their release. According to Education officials, the decision to provide such information is generally based on several considerations, including (1) the quality of the data, (2) Education's confidence that the preliminary findings will be consistent with the final results, (3) the importance and usefulness of the information to policymakers, and (4) in some cases, feedback from IRP and TWG members.

Design Features Influenced Each Study's Ability to Draw Conclusions

For both the LESCP and Prospects studies, key design features influence the studies' ability to draw conclusions on Title I, school reform, or instructional practice. The design of each study has both strengths and potential limitations. For example, Prospects' large, representative sample supported fairly strong conclusions about the effect of Title I, although the study design did not allow researchers to definitively measure how students would have performed if they did not have access to Title I services. The LESCP study is not yet complete, but the depth of information provided by its longitudinal focus is likely to be a key strength. However, limitations to the LESCP study—particularly the study's small, nonrepresentative sample—will restrict researchers' ability to draw strong conclusions from the data, especially about Title I or standards-based reform.

Large, Representative Sample Supported Prospects Study's Ability to Draw Strong Conclusions About Title I

The Prospects study was one of the largest and—at a total cost of $28.8 million—most expensive studies Education has conducted on the Title I program. Prospects data were gathered on a large, nationally representative sample of 372 schools from 1991 to 1994. Students in these schools—including some students that moved—were followed for 3 to 4 years. The Prospects study also collected information on a large number of important factors at the student, family, teacher, school, and district levels. These factors included student and family characteristics, characteristics of the school, teacher standards and expectations, and instructional practices. (See app. I for more information.)

Its rigorous and comprehensive design allowed the Prospects study to support fairly strong conclusions about the effect of the Title I program on student achievement. For example, because Prospects collected data from a large, nationally representative sample, researchers could draw stronger conclusions about the effect of Title I services nationwide than would have been possible with a smaller, nonrepresentative sample. The Prospects data provided information on how Title I services affected student achievement over a 3-year period, rather than providing only a "snapshot" of the program's effect. In addition, the Prospects study captured the experiences and outcomes of highly mobile students—an often underserved population that is at risk for low achievement and dropping out of school.

Prospects' extensive and detailed information at the district, school, teacher, parent, and student levels created a rich database on students

across the country, allowing researchers to explore a variety of factors associated with student educational outcomes. For example, the Prospects study found that the poverty level of the school (over and above the economic status of an individual student) was negatively related to standardized achievement scores and that mobility also had a negative effect on students' academic performance. The researchers also explored relationships between teachers' instructional practices and student outcomes. The study found that an emphasis on comprehension and the development of writing skills were generally positively related to student achievement in reading, and an emphasis on remedial instruction was positively related to math test scores in first through third grade.

Although its comprehensive design clearly strengthened the final study, several design choices limited the conclusions that could be drawn from Prospects' data.

- Prospects' large sample size, although clearly a major strength, led researchers to rely on self-reported survey data in several areas, including instructional practice. Because respondents tend to select socially desirable choices and to overestimate their own progress, the quality of these data may be questionable.
- The Prospects study concentrated largely on students in the elementary grades. While elementary school students constitute the majority of those served by Title I, the program also extends to middle and high school students. Prospects' data provided less information about these students' progress under Title I.[5]
- The Prospects study attempted to compare Title I students with similar students that did not receive Title I services.[6] However, because Title I students were selected by principals and teachers who believed these students most needed extra help, the students who received Title I services may have been more educationally disadvantaged than their non-Title I counterparts. Because students were selected for Title I

[5]Prospects' research design included a seventh grade cohort that was followed through the ninth grade, but limited analysis was done on this group. Title I participation by these students was also lower than for elementary school students.

[6]To construct this comparison group, Prospects' researchers selected those first grade students who attend schools with over 50 percent of students from low income families and who also scored in the bottom quartile of the national standardized test. Within this selected group, the researchers compared Title I students with those students who did not receive Title I services. The Prospects study did not construct similar comparison groups for the third and seventh grade cohorts.

rather than being randomly assigned, the study design did not allow researchers to definitively measure how the same students would have performed in the absence of the program.

- The Prospects study also did not develop clear criteria to measure the effectiveness of the Title I program. As a result, different individuals have interpreted Prospects' findings differently in terms of the success or failure of Title I. Some individuals have interpreted Prospects' results as evidence that the Title I program may not be effective because Title I students did not rise to the achievement level of their peers. Others disagree, saying that without Title I, gaps between Title I students and their peers would have grown over time rather than remain constant.
- Prospects did not provide information on the cost of Title I services in conjunction with their effect on student achievement, limiting policymakers' ability to determine whether the program is a good investment of federal dollars.

Another key limitation of Prospects came about not because of the study design, but because major changes to the Title I program were implemented after the data were gathered in 1991 through 1994 and before the final Prospects report was issued. Consequently, it is difficult to determine how to apply Prospects' results to the redesigned Title I program. For example, at the time the Prospects data were collected, relatively few Title I schools were operating schoolwide programs. By 1999, however, nearly half of Title I schools were operating schoolwide programs. Because schoolwide programs may differ from targeted assistance schools in the way they deliver services, the results from the Prospects study may be less applicable to the Title I program as it is currently structured.

LESCP Likely to Have Limited Ability to Draw General Conclusions, Especially About Title I or Standards-Based Reform

The ongoing LESCP study is smaller and (at a total cost of about $9 million) considerably less costly than the Prospects study. Between 1997 and 1999, LESCP gathered data on a small, nonrepresentative sample of 71 Title I schools from 18 districts in 7 states. LESCP measured student achievement using a national standardized test, similar to the test used in Prospects; however, LESCP also collected student achievement data as measured by state assessments. Like Prospects, the LESCP study also collected information on a large number of factors at the student, family, teacher, school, and district levels. These factors included student and family characteristics, characteristics of the school, teacher professional development, and instructional practices. (See app. I for more detailed information on LESCP's study design).

Because the LESCP study is ongoing, complete information is not available to fully assess the study's strengths and limitations. However, as was true for Prospects, LESCP's design features will contribute to the strengths and limitations of the study and directly affect the conclusions that can be drawn from the data. For example, one strength is that LESCP's data will describe the study schools over a period of several years. In addition, because LESCP gathered detailed data at the district, school, teacher, parent, and student levels, the study can be expected to provide a depth of information on a variety of issues facing these high-poverty schools.

Despite these important strengths, the LESCP study is unlikely to fully satisfy any of its three potential purposes—evaluation of Title I, standards-based reform, or instructional practices—as envisioned by Education, contractors, and panel members. LESCP's major design limitations include the following:

- Most important, LESCP results cannot be generalized beyond the small, nonrepresentative sample of schools, districts, and states used in the study. The small sample size may also make it difficult to compare study results across special populations of students attending Title I schools, such as students with disabilities or students with limited proficiency in English.
- Because district officials chose the schools that participated in the study, the sample is likely to include a greater percentage of higher-achieving schools compared with the population as a whole.
- LESCP's study design, like that of Prospects, relied primarily on self-reported survey data to provide information on important variables such as the implementation of standards-based reform policies and teachers' instructional practices. Because respondents tend to select socially desirable choices and to overestimate their own progress, the quality of these data may be questionable.
- Because LESCP primarily focused on schools rather than students, the study did not follow students who changed schools. Since mobility is associated with lower student achievement, the absence of students who moved is likely to result in overestimates of student achievement in the sample as a whole. Furthermore, this effect may be stronger for some schools in the sample than for others.

In addition to the broad limitations described above, several LESCP design features will restrict the study's ability to draw conclusions about the effects of standards-based reform and Title I, which are important areas for Congress to consider in the reauthorization of Title I. However, these

additional limitations of the study design are likely to have a less serious effect on potential conclusions about instructional practices. Additional limitations include the following.

- No consensus has emerged on how to define and measure standards-based reform. Furthermore, state and local implementation of standards-based reform varies in ways that can be difficult to measure consistently across states and districts.[7] For example, state standards vary considerably in content and level of detail. LESCP measured teachers' ratings of how familiar they were with state standards and assessments, but by design such measures do not capture the content of the standards themselves. The challenges in measuring these practices will make it more difficult to interpret LESCP's results.
- Changes over time in the implementation of standards-based reform policies may also limit LESCP's ability to tie specific reform policies to student achievement. The seven states in the sample were selected to include five states that had already implemented policies related to standards-based reform and two that had not yet done so. However, by the time most of the data were collected, the states in the sample that had been slower to implement these policies had largely caught up with the others in the sample. As a result, there was limited variation in standards-based reform across the seven states, making it more difficult to link this reform to student achievement.
- Because LESCP's sample included only schools that receive Title I dollars, the study design will not allow researchers to compare Title I schools with schools that did not receive Title I funding. Without this comparison, it will be difficult to draw conclusions about the program's overall effectiveness.
- Although Title I services are provided to students in preschool through high school, LESCP data extended only to students in the third through fifth grades, providing an incomplete picture of the effect of Title I services.
- The large majority of schools (58 of 71) in the LESCP sample operated schoolwide programs, and these schools generally had higher poverty rates than the 13 targeted assistance schools in the sample. With such a small number of schools, LESCP will not be able to address a major gap in existing Title I research by drawing conclusions about how the

[7]For more information about this variation, see *Title I Program: Stronger Accountability Needed for Performance of Disadvantaged Students* (GAO/HEHS-00-89, June 1, 2000).

performance of schoolwide programs compares to targeted assistance schools with similar poverty rates.

- Like Prospects, LESCP did not include information on the cost of Title I services in conjunction with their effect on student achievement, thus limiting policymakers' ability to determine whether the program is a good investment of federal dollars.

In addition to the limitations imposed by the study design, LESCP's ability to provide information on the effectiveness of Title I may be complicated by changes in the program since the 1994 reauthorization. Most important, growth in the number of schoolwide programs has complicated efforts to determine the effect of Title I because, under the schoolwide approach, it is more difficult to distinguish the effect of Title I services from the effect of the school's overall instructional program. Consequently, using LESCP data to assess the effect of Title I funds would be even more difficult than such an analysis was for the Prospects study.

In discussing the limitations of the LESCP study, Education staff, panel members, and contractors pointed to two factors that they believe made this study especially challenging to design and implement. First, they pointed out that, over the life of the project, the budget for the LESCP study will be less than one-third of what was spent on Prospects, even without adjusting for inflation. With additional funding, Education staff and panel members told us, LESCP could have included a larger, more representative sample, a wider span of grades, and more qualitative and quantitative data on instructional practice and the implementation of standards-based reform. Second, some individuals stressed the inflexibility of the contracting process, which requires Education to determine detailed study specifications in advance. For example, several panel members pointed out that Education and its contractors may have only a limited ability to respond to panel members' comments on the study's design because making design changes may require Education to re-open the competitive bidding process.

Conclusions

Both the Prospects and LESCP studies produced valuable information, but no single study can fulfill the diverse information needs of the Congress, Education, the states, and principals and teachers. Moreover, individuals often have different views about which research questions are most important to address. Given this diversity, it is especially important to clarify the purpose of research studies so that the potential users of the research can anticipate how they will be able to use the information. In the

case of Prospects, there was general agreement on the study's purpose, and the completed study provided information that was closely related to that purpose. There was much less agreement, however, on the overall purpose of the LESCP project. As a result, even those individuals closely involved with LESCP did not always know what information to expect from the study. Furthermore, a clear statement of purpose is not sufficient, in and of itself, to guarantee a strong research design—especially in a complex field such as education where many related factors contribute to student achievement. Nonetheless, when the purpose of a study is clear, it becomes easier for researchers to identify and minimize the effect of important limitations in the study design. This in turn will enable them to draw stronger conclusions.

Agency Comments and Our Evaluation

Education provided comments on this report, which appear in app. II. In the Department's comments, the Deputy Secretary of Education stated that we accurately described the process used to conduct both the Prospects and LESCP studies. Concerning our finding that the purpose of LESCP was not clearly understood by those involved, Education said that the study was designed to examine both the effects of standards-driven reforms and instructional practices in high-poverty schools. Education also stated that these two purposes are not inconsistent with one another. We agree that a study can have more than one purpose. However, Education's position on the purposes of LESCP was not clearly articulated by the research questions, nor was this view consistently reflected by Department staff, panel members, and study contractors. Furthermore, as our analysis shows, the major limitations of the study's design seriously restrict the conclusions that can be drawn about standards-based reform on the basis of LESCP data. In its comments, Education also stated that GAO recognized that LESCP's design limitations reflected "the need to gather rich, in-depth data within the constraint of limited Title I funding." We did not analyze the extent to which LESCP's funding level influenced the study design; rather, we merely reported that Education identified the funding level as a factor that affected the study. Education also submitted technical comments, which we incorporated where appropriate.

We are sending copies of this report to the Honorable Richard Riley, Secretary of Education, appropriate congressional committees, and other interested parties. We will also make copies available to others upon request. If you or your staffs have any questions about this report, please

call me at (202) 512-7215. Major contributors to this report included Jeff Appel, Sarah L. Glavin, and Sara L. Schibanoff.

Marnie S. Shaul
Associate Director,
Education, Workforce and
Income Security Issues

Comparison of Prospects and LESCP Studies

	Prospects	LESCP
Time frames		
Data gathered	1991-94	1997-99
Interim report released	July 1993	July 1999
Final report released	April 1997	Not yet released
Study cost		
Estimated study cost	$28.8 million	$9 million
Objectives, findings, methodology, and key limitations		
Study objectives	*As written in final report:* To examine the effects of Title I on student achievement and other school-related educational outcomes. *As described in interviews*: Same as in report.	*As written in interim report*: To analyze variation and changes over time in students' performance and teachers' instructional practices in a set of high poverty elementary schools. *As described in interviews*: To analyze the effectiveness of standards-based reform and to evaluate the effectiveness of Title I.
• Major findings	• Students who received Title I services began below the achievement level of their peers. This initial gap remained essentially unchanged as students moved into higher grades. • Characteristics of students and their families account for the largest part of overall variation in student achievement. • The poverty level of the school (over and above the economic status of an individual student) was negatively related to standardized achievement scores. • Academic standards varied between high- and low-poverty schools. • An emphasis on comprehension and the development of writing skills was generally positively related to student achievement in reading. An emphasis on remedial instruction was positively related to math test scores in grades 1-3.	• Some instructional practices, such as frequent repetition, were associated with improved student performance in reading and mathematics. • There was little change in these instructional practices between 1997 and 1998. • Teacher participation in professional development was only modestly associated with differences in classroom practices across teachers. Furthermore, individual teachers did not appear to change their practices as a result of professional development activities.
Data collection strategies	Teacher surveys, standardized tests, student and parent surveys, principal surveys, and a district coordinator survey.	Teacher surveys, standardized tests, state and local assessments, interviews with principals and district Title I coordinators, classroom observations, focus groups of teachers and parents, review of policies and programs of family involvement.

(Continued From Previous Page)

	Prospects	LESCP
Key limitations	• The quality of self-reported data may be questionable. • Information on students in grades above the sixth grade was limited. • Students in comparison groups may have different educational needs than Title I students, which limited the study's ability to draw conclusions. • The study lacked clear criteria to measure the effectiveness of the Title I program. • Major changes occurred in the Title I program before the final Prospects report was issued. • No cost information was given.	• The small, nonrepresentative sample limited the study's ability to draw general conclusions. • Because district officials chose the schools that participated in the study, the sample was likely to include a greater percentage of higher achieving schools compared with the population as a whole. • The small sample size may make it difficult to compare study results across special populations of students attending Title I schools, such as students with disabilities or students with limited English proficiency. • Students that moved were not followed. • Information was provided only for students in grades 3 to 5. • The quality of self-reported data may be questionable. • The difficulty defining and measuring standards-based reform makes it more difficult to interpret the study's results. • The status of reform efforts is changing over time, making it difficult to link standards-based reform to student achievement. • Title I schools could not be compared with non-Title I schools. • Under the schoolwide approach, it is more difficult to distinguish the effect of Title I services from the effect of the school's overall instructional program, which creates difficulties in assessing the impact of Title I. • Results for schoolwide and targeted assistance schools could not be compared. • No cost information was given.
Measures used		
Student achievement	National standardized test—Comprehensive Test of Basic Skills (CTBS).	National standardized test—Stanford 9—and state and local assessments.
Other outcomes	Student attendance and teacher ratings of student performance.	None.
Instructional practices	Frequency of use of various instructional strategies (such as using computers and leading discussion groups) and curricular areas (such as whole number operations, fractions, and vocabulary).	Frequency, duration, and emphasis on various instructional strategies (such as reading aloud, working at a computer, or using calculators) and curricular areas (such as vocabulary, phonics, or multidigit multiplication).

(Continued From Previous Page)

	Prospects	LESCP
Other major variables	Student information on grade retention; preschool attendance; teacher ratings of student ability and motivation; family characteristics; teacher background, certification, experience; estimates of school climate; and school and district characteristics.	Teachers' self-reported familiarity with policy instruments (such as content standards and curricular frameworks); teachers' reports of the amount, type, content, and usefulness of professional development activities.
Sample characteristics		
Sample selection method	Selection was random to ensure a nationally representative sample; however, the 12 largest districts were selected with certainty.	The selection was purposive and stratified by the implementation of standards-based reform in the states and districts.
Number of states	Not available	7
Number of districts	120	18
Number of schools[a]	372	71
Number of students[b]	28,369	2,323
Schools in sample that received Title I funding (%)	66	100
Schoolwide schools in the sample (%)	Not available	81
Targeted assistance schools in the sample (%)	Not available	19
Average poverty rate for schools in the sample (%)	50	74
School sample with poverty rates over 75% (%)	25	35
Urbanicity of school sample (%)	49	Not available

[a]For the Prospects study, this number includes only the original longitudinal cohort.

[b]The number of students is the largest number in the 1-year longitudinal cohort.

Comments From the Department of Education

UNITED STATES DEPARTMENT OF EDUCATION

THE DEPUTY SECRETARY

August 17, 2000

Ms. Marnie S. Shaul
Associate Director
Education, Workforce, and Income Security Issues
General Accounting Office
Washington, DC 20548

Dear Ms. Shaul:

We appreciate the thoroughness with which GAO undertook its review of the information available on the Longitudinal Evaluation of School Change and Performance (LESCP) and the Prospects study. Both of these studies were carried out by our Planning and Evaluation Service (PES).

We believe that GAO accurately describes the open evaluation process both studies used, including contracting out data collection and analysis, soliciting and incorporating comments from expert panels of reviewers, and releasing the results. GAO also notes that both LESCP and Prospects "produced useful results," although the full benefits from LESCP will not be known until study completion this fall.

The GAO report raises an important issue regarding the purpose of LESCP. GAO notes that the LESCP study purpose was not as clearly agreed upon as that of the Prospects study. We want to clarify that a study as complex as LESCP has more than one purpose. In particular, this study was designed to examine the impacts of standards-driven reforms as well as the impacts of instructional practices in high-poverty schools. These two issues are not inconsistent with one another, since a study of the impacts of standards-driven reforms in schools would necessarily take into account, and estimate, differences in impacts due to instructional practices.

An interesting GAO observation is that the LESCP panels were composed mainly of researchers and this contrasted with the greater reliance in Prospects on inclusion of educators as panel members. PES is taking steps to ensure appropriate involvement of educators and other end-users in future evaluation designs. The close cooperation among PES, our program offices, and their constituencies on the evaluation design is especially important to ensure practitioner involvement.

The GAO findings point out potential limitations of the smaller purposive non-representative sample of schools in LESCP. GAO also recognizes that this reflects the

400 MARYLAND AVE. S.W., WASHINGTON, D.C. 20202-0500

Our mission is to ensure equal access to education and to promote educational excellence throughout the Nation.

need to gather rich, in-depth data within the constraint of limited Title I funding. We note that:

- A study of the impacts of standards-based reform in high-poverty schools should only look at those states and districts in which such reforms had taken place. A nationally representative sample of schools would be an inefficient way to identify schools located in districts and states where standards-based reform had taken place.

- Determining the impacts of specific instructional practices in high-poverty schools does not require a national sample. For example, rigorous research on the impacts of instructional practices in reading and mathematics has primarily come from small, purposive samples.

We thank you for the opportunity to comment on the draft report.

Sincerely,

Frank S. Holleman III

Ordering Information

The first copy of each GAO report is free. Additional copies of reports are $2 each. A check or money order should be made out to the Superintendent of Documents. VISA and MasterCard credit cards are accepted, also.

Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

Orders by mail:
U.S. General Accounting Office
P.O. Box 37050
Washington, DC 20013

Orders by visiting:
Room 1100
700 4th St. NW (corner of 4th and G Sts. NW)
U.S. General Accounting Office
Washington, DC

Orders by phone:
(202) 512-6000
fax: (202) 512-6061
TDD (202) 512-2537

Each day, GAO issues a list of newly available reports and testimony. To receive facsimile copies of the daily list or any list from the past 30 days, please call (202) 512-6000 using a touchtone phone. A recorded menu will provide information on how to obtain these lists.

Orders by Internet:
For information on how to access GAO reports on the Internet, send an e-mail message with "info" in the body to:

info@www.gao.gov

or visit GAO's World Wide Web home page at:

http://www.gao.gov

To Report Fraud, Waste, or Abuse in Federal Programs

Contact one:

- **Web site: http://www.gao.gov/fraudnet/fraudnet.htm**
- **e-mail: fraudnet@gao.gov**
- **1-800-424-5454 (automated answering system)**

PRINTED ON RECYCLED PAPER

United States
General Accounting Office
Washington, D.C. 20548-0001

Bulk Rate
Postage & Fees Paid
GAO
Permit No. G100

Official Business
Penalty for Private Use $300

Address Correction Requested

BEST COPY AVAILABLE